Causal effect of early initiation on adolescent smoking patterns

M. Christopher Auld[1]
University of Calgary

August 2002
This version: December 2002

Abstract

A key concern in policy debates over youth smoking is whether preventing children from smoking will stop them from smoking as adults or merely defer initiation into smoking. This paper estimates smoking status in late adolescence viewing smoking at age 14 as an endogenous "treatment" on subsequent smoking, an approach which disentangles causation from unobserved heterogeneity and allows the model to capture the theoretical prediction that addictiveness varies across individuals. Exploiting large tax changes across time and across regions in Canada in the early 1990s, a structural model of potential outcomes is estimated to recover distributions of heterogeneous addictiveness. The results suggest that smoking is highly addictive for the average youth but less so for youths who actually do initiate early or are likely to be induced to initiate early at the margin. Smoking initiation is found to be price sensitive, but conditional on previous smoking behavior youth smoking is price inelastic.

JEL Classification: I1, C3
Keywords: smoking, addiction, causal model, self–selection

[1] I thank the Alberta Heritage Foundation for Medical Research for financial support.

Causal effect of early initiation on adolescent smoking patterns

1 Introduction.

This paper seeks to answer the question:

> Suppose an individual is induced not to initiate daily smoking by age 14. How will smoking behavior in late adolescence be affected?

Consider a policy which successfully reduces smoking rates among fourteen year olds. If the policy merely defers the initiation of smoking into later adolescence or early adulthood, it will fail to have substantial long–term effects on health. However, if it reduces the number of people who *ever* smoke, it will be successful in reducing eventual smoking–related morbidity and mortality. The causal effect of youth smoking on smoking patterns later in life is, then, of critical policy importance. Using a large survey of in and out of school adolescents, a dynamic structural model of youth smoking behavior is developed to break the observed correlation in smoking patterns over time into components attributable to unobserved heterogeneity in smoking propensity and heterogeneous addiction to tobacco.

It is sometimes assumed in the public health literature that deterring youth smoking must decrease adult smoking, yet this result does not follow from the observed high correlation in smoking status over time. In the primary data used in this paper, for example, early smoking initiators (age<15) are far more likely to report being daily smokers later in adolescence: A youth who reports initiating into smoking by age 14 is 5.5 times more likely to smoke later in adolescence than a youth who did not begin smoking by age 14. The key question is whether this pattern obtains because initiating early *causes* higher smoking rates later in adolescence or rather because the same factors which induce early initation also induce smoking later in adolescence. To the extent that these patterns reflect addiction, preventing initiation will cause lower smoking rates later in life. But if such patterns instead reflect the fact that some teenagers are more likely to smoke for reasons other than addiction, preventing initiation at a young age may simply shift initiation to a later age. Since the health effects of smoking are cumulative over decades and are small if an individual quits at an early age, a policy which merely deters but does not prevent smoking can be expected to have little impact on health over the life cycle. If, conversely, deterring smoking at a young age reduces the probability a person *ever* smokes, policies aimed at preventing youth smoking may be highly effective in reducing smoking–related costs (Glied 2002).

Recently, a controversy has arisen in the econometrics literature on the effect of price changes on youth smoking behavior. Several studies seem to confirm that not only only is youth smoking affected significantly by price, youth smoking is considerably more price elastic than smoking by adults (Lewitt *et al.* 1981, USDHHS 1994, Evans and Huang 1998, Harris and Chan 1999, Tauras and Chaloupka 1999). But DeCicca, Kenkel, and Mathios (1999) report that initiation into smoking at ages 13 through 16 is not affected by price, and a similar result is discovered by Gruber (2000). Jones and Forster (2001) find age of initiation is little affected by price but that price increases the hazard out of smoking. Tauras *et al* (2001) report that probability of initiation is significantly reduced by higher prices. Glied (2002) focuses on the long-term effects of prices faced at age 14. Glied finds that taxes at age 14 have substantial effects on contemporaneous smoking, but that the effect diminishes rapidly over time.

Gilleskie and Strumpf (2000) emphasize that much of this literature is difficult to interpret because past smoking behavior is not modeled. What is needed is a structural model. Gilleskie and Strumpf (2000) disentangle the effects of unobserved heterogeneity and addiction using an estimable stochastic dynamic decision making approach. They find important differences in the elasticities of never smokers and previous smokers, with previous smokers exhibiting essentially no price sensitivity.

A problem identified in the literature using U.S. data is the endogeneity of State–level tobacco taxes. States with high taxes are likely to be States with high levels of anti-smoking sentiment, which introduces correlation between unobserved determinants of smoking and tax rates, that is, tax rates may not be valid instruments for smoking (DeCicca, Kenkel, and Mathios 1999). If these level changes are differenced out using fixed effects, remaining variation is small, so taxes are then weak instruments (Tauras *et al* 2001). In this paper, transient tax changes in some of Canada between 1991 and 1994 are exploited to reduce these problems. The primary data are a subsample aged 15 to 19 drawn from the Canadian 1994 Youth Smoking Survey. This survey was conducted in late 1994, so respondents became 14 sometime between 1989 and 1994, encompassing the tax changes.

The econometric model treats early initiation as a non–randomly assigned "treatment" on the binary outcome "smokes in late adolescence." A structural model of potential smoking probabilities is estimated. The model allows the determinants of initiation, smoking if initiation was early, and smoking if initiation was not early to vary arbitrarily, and a flexible error structure allows explicit modeling of selection on unobservables. From the estimated model, causal effect distributions may be evaluated, which in this context may be interpreted as distributions of levels of smoking addictiveness. Thus, the model is able to capture Becker and Murphy's (1988) emphasis on the fact that the addiction is a behavioral trait which varies across individuals.

2 Causal model for adolescent smoking.

Because of its addictive nature, smoking status at any time depends in complex manner on the entire sequence of previous starts and quits, and thus the entire sequence of past and future prices. Recasting this complex behavior as a two–step sequential problem allows the core issue — how much of the correlation between current and past smoking behavior is causal and how much reflects unobserved heterogeneity — to be addressed in a feasible manner. Smoking at age 14 is modeled as a non-randomly assigned "treatment." Quasi-randomization occurs in the form of a natural experiment: Variation in prices at age 14 affect smoking decisions at age 14, but do not affect subsequent smoking decisions conditional on decisions at age 14.

Cast in this manner, smoking decisions over time can be modeled using the well–developed framework for analysis of treatment effects beginning with Neyman (1923) and recently developed in numerous papers including Heckman and Robb (1985), Imbens and Angrist (1994), Angrist, Imbens, and Rubin (1996), and Heckman (1997).[1] The analysis here draws on that of Aakvik, Heckman, and Vytlacil (2000), who develop a model appropriate for binary outcomes when responses to treatment are heterogeneous. The model is extended to the case of maximum likelihood estimation of the system when the errors follow a multivariate Student–t distribution. Heckman, Tobias, and Vytlacil (2000) discuss two-step estimation of the model with Student errors and continuous outcomes. Similar causal models are studied in Imbens and Rubin (1997), Aakvik *et al* (2000), and Chib and Hamilton (2001).

Adopting the notation common to the causal effects literature, denote the endogenous outcomes

D indicator for early initiation into smoking
Y_1 indicator for smoking in late adolescence if early initiator
Y_0 indicator for smoking in late adolescence if not early initiator.

[1] See also the critique of this approach presented by Ferrall (2001).

The inference problem arises primarily because Y_{1-j} is not observed if Y_j is observed, $j = 0, 1$. The econometric strategy is to estimate a structural model capable of generating counterfactual smoking probabilities, taking into account selection into early initiation on both observables and unobservables. The structural model is written in terms of latent outcomes. Smoking decisions are determined by

$$\begin{aligned} D^* &= Z\gamma + U_D \\ Y_1^* &= X\beta_1 + U_1 \\ Y_0^* &= X\beta_0 + U_0, \end{aligned} \tag{1}$$

where X, Z are vectors of exogenous covariates, $U = (U_D, U_0, U_1)$ are is a disturbance vector, and $(\gamma, \beta_0, \beta_1)$ are vectors of parameters to be estimated. Latent outcomes are related to observed outcomes according to $D = 1[D^* > 0], Y_1 = 1[Y_1^* > 0], Y_0 = 1[Y_0^* > 0]$.

2.1 Error structure.

The errors are assumed to follow a multivariate Student's–t distribution with degrees of freedom ν. This distribution has the advantages that it nests the Gaussian distribution which obtains as $\nu \to \infty$, but exhibits "thicker tails" than that of the Gaussian distribution for low values of ν. The error density is then

$$t_\nu(u, \Sigma) = \frac{\Gamma(1/2(\nu + 3)}{\Gamma(\nu/2)\sqrt{|\Sigma|(\nu\pi)^3}} \left[1 + \left(\frac{1}{\nu}\right) u'\Sigma u\right]^{-(1/2)(\nu+3)} \tag{2}$$

where

$$\Sigma = \begin{pmatrix} 1 & \sigma_{D1} & \sigma_{D0} \\ \sigma_{D1} & 1 & 0 \\ \sigma_{D0} & 0 & 1 \end{pmatrix}. \tag{3}$$

Σ is the covariance matrix of the errors when $\nu > 2$. Notice the covariance between U_1 and U_0 has been normalized to zero, because this parameter is not identified by the data. The diagonal elements are normalized to unity because scale is not identified for any equation.

2.2 Estimation.

The estimation method employed is full information maximum likelihood, such that asymptotic efficiency is obtained and the inherent nonlinearities in the model are captured. Estimation proceeds by maximizing the log of the likelihood over the parameters $\theta = \{\gamma, \beta_0, \beta_1, \sigma_{D0}, \sigma_{D1}\}$. Algorithms due to Genz and Bentz (2002) were used to evaluate rectangle probabilities under the multivariate Student's-t density. A combination of a gradient-free simplex method and Newton–Raphson based methods was used to converge to the maximum. See Auld (2002) for derivation of the likelihood and other details.

3 Data.

The primary dataset is the 1994 Youth Smoking Survey (YSS). The YSS was conducted by Statistics Canada in Fall 1994 to gather information on youth smoking behavior. This paper uses the sample of 15 to 19 year olds drawn as a supplement to the Labor Force Survey. A key advantage of these data is the sampling universe includes youths who have dropped out of school, as Gilleskie and Strumpf (2000) find, using a similarly rich dataset, that smoking behavior of dropouts and students differ. Table 1 displays definitions of variables used in the analysis and descriptive statistics.

Early initiation status is ascertained from retrospective questions. A respondent is classified as an early initiator if they responded that they had smoked a whole cigarette, and had smoked at least one whole cigarette every day for seven consecutive days, and they were 14 or younger when they first began such smoking behavior. A respondent is classified as a late adolescent smoker if they smoked on least 21 days in the last month. Roughly 92% of respondents who were classified as late adolescent smokers reported they smoked every day during the preceding month.

Socioeconomic controls include dummies for year of birth, self-reported health status, whether the respondent is still in school and age they left out if they are not, and whether they consider themselves good, average or poor students if they are still in school. Whether or not the respondent is still in school, they are classified as having a job or not, and the hours per week they work is recorded if they are employed.

Peer effects on smoking are frequently emphasized in both the economics and public health literatures (Ary and Biglan 1988, Krauth 2001). Controls for peer effects on smoking behavior include indicators for parents' current and past smoking status, number of siblings or other non–parental household members who smoke, and number of close friends who do *and* *do not* smoke. Further, a variable is constructed indicating the respondent reported that "more than half" or "almost all" the teachers at their school smoke.

Finally, measures of perceived non-pecuniary costs of smoking were constructed from the YSS. Respondents were asked a sequence of questions regarding whether various specific ailments could result if "someone smoked for many years," such as lung cancer, heart disease, bronchitis, and so forth. The number of "yes" responses to these questions is included as a measure of subjective health risks, as is an indicator the student was taught in school that smoking poses health risks. If the respondent has a job, an indicator that smoking is restricted at the workplace is included. Finally, an indicator is constructed for the student being able to correctly state the minimum legal age to purchase tobacco in their Province of residence, and the interaction of this dummy with a dummy indicating the Provincial legal smoking age is 19.

Notice that most of these variables are recorded in late adolescence; the YSS does not contain repeated observations on sociodemographic information. However, following an argument of DeCicca, Kenkel, and Mathios (2002), these responses are included as covariates in the early initiation equation as well. Heterogeneity in preferences, influences, school performance, health, and family structure in late adolescence is likely to reflect similar heterogeneity at age 14, and indeed we will find that many of the covariates as reported in late adolescent are highly significant determinants of early initiation.

Province-specific tobacco product price indexes, deflated by the all-items CPI, as reported by Statistics Canada are merged with the YSS data. Price in September 1994 is used as the relevant price in late adolescence (denoted **price1**), and price in September of the year the respondent turned 14 is used to proxy price at age 14 (denoted **price0**) . Use of one-year moving averages of monthly prices instead of these measures produced qualitatively similar results in preliminary estimation, but with less precisely measured demand slopes. Figure 1 displays the price data, showing substantial variation across time and provinces.

4 Results.

The structural model consists of initiation and state-dependent smoking propensity equations as described by display (1). Price at age 14 is included in the initiation equation and excluded from the late adolescent smoking equations. The model then makes use of the intuition that, conditional on past consumption decisions, past prices only affect current consumption through intertemporal complementarity or substitutability, that is, addiction (Chaloupka 1991). Formally, future price should be included in the initiation equation, but, as in Gilleskie and Strumpf (2000), current and future price were found to be highly colinear which made estimates difficult to interpret while adding little explanatory power to the model. This result is also consistent with the explanation that youths were unable to make good forecasts of prices years into the future at the time of initiation.

4.1 Model selection.

Model selection proceeded by repeatedly estimating the structural model as the degrees of freedom for the multivariate Student's-t distribution was varied. The results of this exercise are presented in Table 2. For these data, the textbook Gaussian model is massively rejected against "fatter tailed" densities. The preferred model is Student with five degrees of freedom.[2] The standard errors for the structural parameter estimates need not be corrected for this search over degrees of freedom because the Fisher information in the model is block diagonal with respect to degrees of freedom (Lange, Little, and Taylor 1989).

4.2 Structural parameters estimates.

Structural parameter estimates for the preferred model are displayed in Table 3. Price at age 14 is a statistically significant predictor of early initiation into smoking (t=-2.11). In single-equation probits of smoking in late adolescence, price had no significant effect after stratifying by early initiation status (t=0.41 for early initiators, t=-0.36 for non-early initiators). Thus, price at age 14 satisfies the conditions required for an instrumental variable: It affects the treatment decision but does not affect outcomes conditional on treatment. Price has smaller effects in late adolescence conditional on early

[2] Attempts to fit the model with less than five degrees of freedom yielded erratic results, such as pseudo-covariance matrix estimates pushing up against the positive definiteness constraint and predominately negative causal effects. Geweke (1993) warns against using Student distributions with less than five degrees of freedom.

smoking behavior. Table 6 shows marginal elasticity estimates.[3] Table 6 shows that initiation is price elastic, but, conditional on early initiation behavior, demand is inelastic for non–early initiators and essentially unrelated to price for early initiators.

Comparing determinants of smoking for early initiators and non-early initiators shows that generally smoking behavior of the former group is less sensitive to changes in smoking determinants. That is, not only price but most determinants of smoking have less effect for addicted youths. Models which do not allow smoking determinants to vary with previous smoking behavior may assume away economically meaningful changes in incentives.

Estimated effects of the other covariates are typically of the anticipated sign and are precisely estimated. Cohort effects are large in the early initiation equation, with older cohorts less likely to smoke all else equal. In the late adolescent smoking equations the birth year dummies capture both the effect of age on smoking propensity and cohort effects, which cannot be disentangled without strong cross-equation restrictions.

Respondents who report they are still in school in 1994 were less likely to be early initiators, except if they report being poor students, in which case they were about as likely as school leavers to initiate early (note the omitted category is school leavers). Males and females initiate early at roughly the same rate, but males are more likely to take up smoking in

[3]Marginal elasticity η calculated as the mean over individual–specific arc elasticities with respect to a 10% price change, $\eta = N^{-1} \sum_{t=1}^{N} [10(T' - T)/T]$, where T is a predicted smoking probability and T' the predicted probability prices has been increased 10%.

late adolescence regardless of whether they initiate early. Respondents who smoke are likely to report diminished health status; whether this is because smoking causes substantially lower perceived health for youths or because unhealthy youths are more likely to take up smoking cannot be inferred from these data. Youths who believe smoking is more damaging to health are *more* likely to have initiated early and more likely to take up smoking later in adolescence if they did not initiate early. This result is consistent with the findings of Viscusi (1992) and Agee and Crocker (2001).

The most striking result in Table 3 is that the smoking behavior of friends and family is a far more powerful predictor of smoking than prices, risks, and other demographic characteristics. We may not, however, infer that having friends and family who smoke causes the respondent's smoking behavior due to endogeneity of several varieties discussed in, for example, Manski (2000). However, the mechanisms through which parents' smoking are endogenous are weaker than for friends, and particularly mother's smoking behavior is strongly correlated with respondent's smoking behavior. Further, conditioning on the potentially endogenous smoking behavior of friends is likely to result in better estimates than ignoring this information when past prices instrument past smoking behavior, as past prices affect current smoking through both addiction and the behavior of friends.[4]

Finally, note that σ_{D0} and σ_{D1}, the estimated correlations between the unobservables, are large. The signs suggest that youths who are likely to

[4]Since past prices affect current behavior through the effect they have on the smoking behavior of others, the stable unit treatment treatment value assumption (SUTVA) of the causal effects model is violated if social interactions are not held constant.

initiate early also tend to initiate late, that is, that unobserved heterogeneity drives much of the intertemporal correlation in smoking status. Conditional on observed characteristics, increasing U_D, propensity to initiate early, increases probability of smoking late if early initiation does not occur and decreases probability of smoking late if early initiation does occur.

4.3 Heterogeneous addiction.

This section discusses the causal effect of early initiation on late adolescent smoking probability. The causal effects of interest are means of $Y_1 - Y_0$ over different conditioning sets, as discussed in for example Aakvik, Heckman, and Vytlacil (2000). Auld (2002) presents the expressions used to evalate the causal effect parameters for the multivariate Student model. In the present case, the causal effect may be considered to measure the level of addictiveness of smoking in the sense of Becker and Murphy (1988), who define an activity as addictive if increases in current consumption *cause* increases in future consumption. An advantage of adopting the potential outcomes framework is addictiveness can be allowed to vary with respondents' observed and unobserved characteristics. If instead conventional econometric models of the form

$$Y_1 = X\beta + \delta Y_0 + \text{noise} \tag{4}$$

were estimated, the parameter δ generally does not converge to any economically meaningful value even if Y_0 is instrumented and the instruments are valid. Instead, it converges to a weighted average of causal effects, where the weights depend on correlations between the unobservables (Angrist and Imbens 1994). Thus, the potential outcomes framework both avoids a statistical difficulty and provides insight into the theoretical prediction that addiction to a given commodity should vary across individuals.

Causal effects are summarized by the kernel density estimates displayed in Figure 2 and the mean effects listed in Table 4.[5] Consider first the crudest estimate of the causal effect of early initiation on late adolescent smoking. The difference in sample means is

$$E[Y_1 - Y_0] = 0.797 - 0.144 = 0.653. \tag{5}$$

This effect confounds the causal effect of interest with selection on both observables and unobservables. The average treatment effect $ATE(X) = E[Y_1 - Y_0|X]$ gives the causal effect conditional on covariates averaged with respect to the distribution of U^D. It measures how a randomly selected youth would respond to being forced to initiate early or prevented from initiating early. The mean ATE is slightly larger than the raw difference in sample means. Table 4 and Figure 2 show its probability mass is centered on large causal effects. However, it is also not an economically interesting estimate, as it measures effects mostly for youths unlikely to initiate early regardless of policy.

The effect of treatment on the treated $TT(X) = E[Y_1 - Y_0|X, D = 1]$ measures the effect of early initiation on late adolescent smoking in the sample of youths who actually did initiate early. It is much smaller than

[5] Standard errors and 90% confidence intervals for the mean effects parameters calculated using a parametric bootstrap off the estimated asymptotic distribution of the structural parameters. 1,000 bootstrap replications were used.

the average treatment effect with a mean of 0.104 and the 90% confidence interval for the mean effect contains zero. The selected sample of youths who choose to initiate early experience less addiction than a randomly selected youth would.

The marginal treatment effect is the effect of early initiation evaluated at the point where respondents are indifferent over whether to initiate early. In this model, it is also the local instrumental variables parameter of Heckman and Vytlacil (2000). $MTE(X) = E[Y_1 - Y_0|X, U_D = -Z\gamma]$ measures the causal effect for individuals induced into or out of early initiation by a marginal change in Z. The marginal treatment effect is on average larger than the effect of treatment on the treated and smaller than the average treatment effect.

Table 5 illustrates heterogeneous response across observed characteristics. Table 5 shows the results of ordinary least squares regression of $MTE(X)$ on X.[6] Interpreting the marginal treatment effect as the level of addictiveness of smoking for individuals indifferent over early initiation decisions, we see the effect of increased "addictive capital" at age 14 diminishes over time, with 19 year olds 30 percentage points less affected by early initiation than 15 year olds. This result is consistent with a rational addiction framework where past consumption affects current addiction less as past and current consumption become farther apart in calendar time. We also see smoking is less addictive for youths whose friends, family and teachers smoke

[6]Note the standard errors are uncorrected; if this were a simple exogenously switching linear model the fit would be perfect.

even though these youths are most likely to smoke. Youths most likely to smoke tend to be youths for whom smoking is less addictive, in the sense that their current smoking decisions are less affected by their past smoking decisions.

Finally, consider the local average treatment effect, $LATE(X) = E[Y_1 - Y_0|X, u_D \in (z, z')]$, where $z = Z\gamma$ and $z = Z'\gamma$ and Z' differs from Z in that the value of the instruments has been varied. LATE measures the treatment effect in the subpopulation who change their treatment status in response to change in the value of the instruments. Table 5 displays LATE defined with respect to a 75% change in the price of tobacco. Youths deferred from initiating into smoking by such a price increase would exhibit an average 20% lower probability of smoking in late adolescence as a result, holding the price of tobacco in late adolescence constant.

4.4 Simulating counterfactual tax policy.

The cut in tax rates in 1994 has been widely criticized by anti–smoking groups (e.g. Canadian Cancer Society 1999), and Hamilton *et al* (1997) present evidence that quit rates were lower and initiation rates higher in Provinces in which taxes were reduced in 1994. In this section, the estimated structural model is used to conduct an experiment evaluating the effect on youth smoking rates from a change in tobacco taxes: Suppose *neither* the 1994 decrease nor the 1991 increase in taxes occurred. In particular, imagine prices from September 1990 through September 1994 were held constant at their September 1990 levels rather than the levels which occurred historically. How would youth smoking rates in late 1994 have differed? The counterfactual question is difficult to answer within a reduced form empirical approach: The total effect depends on all of: the effect of price changes on initiation decisions, the addictiveness of smoking among youths affected by price changes and how this addictiveness varies with observed and unobserved characteristics, and on the contemporaneous effect of price on smoking propensity, which in turn varies with previous smoking behavior.

Table 7 summarizes the results of this exercise. The first two columns show that the structural model is able to mimic the frequencies observed in the data to two decimal places when evaluated at historical prices. At the counterfactual prices, early initiation rises three percentage points from 13% to 16%. Because the causal effect is relatively low for youths induced to alter behavior by such price changes, the long-term effect is smaller. Counterfactual eventual smoking is predicted to be 1.4 percentage points higher than observed.

The 1.4 percentage point change in eventual smoking resulting from not increasing prices during 1991-1994 may seem inconsequential. However, the cohort who were 15 through 19 in 1994 numbered roughly two million people. If the tax hike had not occurred, about (0.014)(2M)=28,000 more of the members of this cohort would have been daily smokers in late adolescence. If we conservatively assume that only half of those would have continued smoking and only one in four of those would die from a smoking-related illness, we conclude that the transient tax hike will eventually prevent roughly

3,500 smoking-related deaths in this cohort.

5 Conclusions.

This paper considered a dynamic empirical model of youth smoking behavior. Viewing early initiation into smoking as a non-randomly assigned "treatment" and later smoking behavior as a dichotomous outcome, a causal effects model with a robust error structure was estimated. The results suggest that tobacco is less addictive for youths who are likely to smoke than for randomly selected youths, but nonetheless smoking is found to be addictive for virtually all youths. For example, among youths who would change their behavior in response to a 75% change in tobacco prices, forcing early initiation into smoking is predicted to cause an average 20 percentage point increase in probability of smoking in late adolescence. These effects vary with both observed and unobserved determinants of youth smoking. These results suggest that policies which deter early initiation may not have large effects on eventual smoking rates, consistent with the results of Glied (2002).

The estimated structural model suggests that initiation decisions in early adolescence are price sensitive (with an elasticity of -1.3), but that conditional on early smoking status later decisions are little affected by price. The model further reveals that smoking is less sensitive to changes in incentives other than price for addicted youths.

The structural model was used to evaluate a policy counterfactual: What if the Canadian government had not increased tobacco taxes during the period 1991-1994? The results indicate eventual smoking would have been about 1.4 percentage points higher had the tax increase never occurred. A back-of-the-envelope calculation using that result suggests that, had the tax increase never happened, roughly 3,500 more deaths from smoking-related illness would eventually have occurred in this cohort of about two million people.

6 References

Aakvik, A., J. Heckman, and E. Vytlacil (2000) "Treatment effects for discrete outcomes when responses to treatment vary among observationally identical persons: An application to Norwegian Vocational Rehabilitation Programs," forthcoming *Journal of Econometrics.*

Agee, M. and T. Crocker (2001) "Smoking parents' evaluation of their own and childrens' health," manuscript.

Angrist, J., and Imbens, G., (1994), "Identification and estimation of local average treatment effects," *Econometrica* 62:467-75.

Angrist, J., G. Imbens, and D. Rubin (1996) "Identification of causal effects using instrumental variables," *Journal of the American Statistical Association* 91:444-455.

Ary, D. and A. Biglan (1988) "Longitudinal changes in adolescent cigarette smoking behavior: Onset and cessation." *Journal of Behavioral Medicine*, 11(4), 361-382.

Auld, M. (2002) "Robust system estimation of causal effects on binary outcomes with application to effect of alcohol abuse on employment," manuscript.

Canadian Cancer Society, Non–smokers' Rights Association, Physicians for a Smoke–free Canada, and Quebec Coalition for Tobacco Control (1999) "Surveying the Damage: Cut-rate tobacco products and public health in the 1990s." Manuscript.

Becker, G. and K. Murphy (1988) "A theory of rational addiction," *Journal of Political Economy* 96:675-700.

Chaloupka, F. (1991) "Rational addictive behavior and cigarette smoking," *Journal of Political Economy* 99(4):722-42.

Chaloupka, F. and H. Wechsler (1997) "Price, tobacco control policies, and smoking among young adults," *Journal of Health Economics* 16:359-73.

Chib, S. and B. Hamilton (2001) "Bayesian analysis of cross-section and clustered treatment models," *Journal of Econometrics* 97(1):25-50.

DeCicca, P., D. Kenkel, A. Mathios (2002) "Putting out the fires: Will higher taxes reduce the onset of youth smoking?" *Journal of Political Economy* 110:144-169.

Ferrall, C. (2001) "Estimation and inference in social experiments," manuscript, Queen's University at Kingston.

Genz, A. and F. Bretz (2002) "Methods for the computation of multivariate t probabilities," forthcoming, *Journal of Computational and Graphical Statistics.*

Geweke, J. (1993) "Bayesian treatment of the independent student t linear model," *Journal of Applied Econometrics* 8: S19-S40.

Gilleskie, D. and K. Strumpf (2000) "The behavioral dynamics of youth smoking," NBER 7838.

Glied, S. (2002) "Youth tobacco control: Reconciling theory and empirical evidence," *Journal of Health Economics* 21:117-35.

Gruber, J. (2000) "Youth smoking in the U.S.: Prices and policies," NBER 7506.

Gruber, J., A. Sen, and M. Stabile (2002) "Estimating price elasticities when there is smuggling: The sensitivity of smoking to price in Canada," NBER 8962.

Hamilton, V., C. Levington, Y. St-Piere, and F. Grimard (1997) "The effect of tobacco tax cuts on cigarette smoking in Canada." *Canadian Medical Association Journal* 156:187-91.

Harris, J. and S. Chan (1999) "The continuum of addiction: Cigarette smoking in relation to price among Americans aged 15-29," *Health Economics* 8:81-86.

Heckman, J. (1997) "Instrumental variables: A study of implicit behavioral assumptions used in making program evaluations," *Journal of Human Resources* 32:441-462.

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998) "Characterizing selection bias using experimental data," *Econometrica* 66:1017-1098.

Heckman, J. J. Tobias, and E. Vytlacil (2000) "Simple estimators for treatment parameters in a latent variable framework with an application to estimating the returns to schooling," NBER #7950.

Jones, A. and M. Forster (2001) "The role of taxes in starting and quitting smoking: duration analysis of British data," *Journal of the Royal Statistical Society: Series A,* 2001, 164, 517-547.

Krauth, B. (2001) "Peers and prices: Explaining the black/white youth smoking gap," manuscript, Simon Fraser University.

Manski, C. (2000) "Economic analysis of social interactions," NBER 7580.

Neyman, J. (1923) "Statistical problems in agricultural experiments," supplement to *Journal of the Royal Statistical Society* 2:107-80.

Tauras, J. and F. Chaloupka (1999) "Determinants of smoking cessation: An analysis of young adult men and women," NBER 7262.

Viscusi, K. (1992) *Smoking: Making the risky decision,* Oxford: Oxford.

Table 1
Variable definitions and descriptive statistics

variable	definition	mean	std. dev.
smokes0	=1 early smoking initiator	0.130	0.336
smokes1	=1 daily smoker in 1994	0.229	0.420
price0	real tobacco price at age 14	0.921	0.119
price1	real tobacco price 1994	0.748	0.179
born 1978	=1 born in 1978	0.213	0.409
born 1977	=1 born in 1977	0.202	0.402
born 1976	=1 born in 1976	0.180	0.384
born 1975	=1 born in 1975	0.163	0.369
father smokes	=1 father currently smoker	0.341	0.474
father former smoker	=1 father former smoker	0.345	0.475
sibling smokers	# of siblings or other non-parental household member who are smokers	0.175	0.722
mother smokes	=1 mother currently smoker	0.312	0.463
mother former smoker	=1 mother former smoker	0.272	0.445
n.s. friends	# non–smoking close friends	3.943	4.019
smoking friends	# close friends who smoke	2.265	3.927
smoking teachers	=1 one–half or more teachers smoke	0.247	0.431
taught smoking risky	=1 taught in school smoking is unhealthy	0.810	0.391
good student	=1 above average school performance	0.287	0.452
medium student	=1 average school performance	0.530	0.499
bad student	=1 below average school performance	0.032	0.178
has job	=1 has a paying job	0.515	0.499
hours worked / week	# hours works for pay if employed	9.143	13.618
health above avg	=1 health is above average	0.307	0.461
health below avg	=1 health is below average	0.060	0.239
workplace restrictions	=1 smoking restricted at workplace	0.244	0.430
# smoking risks named	# health problems named (max=10)	2.305	1.310
knows lgl. age & age=19	=1 Legal smoking age=19 & R. knows this	0.058	0.233
knows legal smoking age	=1 R. correctly states legal smoking age	0.813	0.389
male	=1 male	0.510	0.499

Note: n=9,139.

Table 2
Model selection

d.f.	log–likelihood
5	-5,625.47
6	-5,635.30
10	-5,663.18
15	-5,682.29
20	-5,693.77
∞	-5,747.76

Notes:
- Each row corresponds to estimates of the structural model conditional on the errors having a multivariate student's-t distribution listed degrees of freedom, where infinite degrees of freedom denotes the Gaussian model.
- All listed distributions may be rejected in favor of t_5 at conventional significance levels ($2[L_5 - L_\nu] \sim \chi^2_1$ under null that true value of ν is 5).

Table 3
Maximum likelihood estimates of structural smoking model.

	D		Y_1		Y_0	
variable	coef	t-ratio	coef	t-ratio	coef	t-ratio
price0	-1.271	-2.114				
price1			0.141	0.386	-0.213	-1.656
born 1978	-0.285	-4.092	0.825	4.170	0.454	2.815
born 1977	-0.509	-6.837	0.680	3.281	0.584	3.011
born 1976	-1.133	-6.992	0.839	2.857	0.611	2.709
born 1975	-1.124	-5.776	0.675	2.008	0.729	3.121
father smokes	0.122	2.012	-0.000	-0.000	0.153	2.515
father former smoker	0.002	0.032	-0.067	-0.451	0.088	1.405
sibling smokers	0.204	7.997	0.239	3.278	0.377	13.324
mother smokes	0.430	7.208	-0.101	-0.507	0.378	6.293
mother former smoker	0.236	3.603	-0.198	-1.057	0.289	4.833
n.s. friends	-0.246	-21.557	-0.060	-0.660	-0.304	-30.049
smoking friends	0.073	13.437	0.007	0.307	0.103	18.789
smoking teachers	0.100	1.750	0.007	0.046	0.040	0.683
taught smoking risky	-0.058	-1.003	-0.023	-0.142	-0.223	-3.842
good student	-0.466	-4.985	-0.176	-0.513	-0.575	-6.585
medium student	-0.306	-3.741	-0.154	-0.524	-0.318	-4.238
bad student	0.049	0.391	-0.413	-1.142	-0.037	-0.272
has job	0.025	0.356	0.094	0.521	-0.041	-0.541
hours worked / week	0.004	1.726	-0.008	-1.280	0.005	2.017
health above avg	-0.298	-4.574	0.004	0.021	-0.470	-7.905
health below avg	0.457	6.280	0.391	1.463	0.709	8.439
workplace restrictions	-0.298	-4.308	0.226	1.108	-0.043	-0.612
# smoking risks named	0.087	4.818	-0.000	-0.004	0.148	8.566
knows lgl. age & age=19	0.130	1.070	0.191	0.498	0.023	0.232
knows legal smoking age	0.398	4.984	0.361	1.469	0.599	7.121
male	-0.013	-0.280	0.266	1.965	0.236	4.388
constant	0.317	0.517	1.134	1.817	-1.543	-4.951
σ_{D1}			-0.590	2.117		
σ_{D0}			0.629	2.974		

Notes: Columns labeled D denote estimates of initiation equation, Y_1 smoking status in late adolescence if early initiator, and Y_0 smoking status if not early initiator. Log-likelihood=-5,625.47.

Table 4
Heterogeneous causal effect estimates.

effect	mean effect	std. dev.	Bootstrapped: std. err.	percentile: 5^{th}	95^{th}
ATE(X)	0.707	0.217	0.086	0.498	0.769
TT(X)	0.104	0.141	0.087	-0.037	0.248
MTE(X)	0.339	0.200	0.138	0.105	0.554
LATE(X)	0.198	0.158	0.128	-0.005	0.408

Notes:
- LATE defined with respect to a 75% increase in cigarette prices.
- "Standard deviation" measures variation of effects across individuals whereas "standard error" and the percentile estimates are estimates of the sampling distribution of the mean effect based on 1,000 bootstrap samples.

Table 5
How addictiveness varies with observables

variable	coefficient	t-ratio
born 1978	0.010	2.61
born 1977	-0.140	-35.47
born 1976	-0.239	-56.02
born 1975	-0.299	-57.92
father smokes	-0.023	-7.01
father former smoker	-0.047	-14.70
sibling smokers	-0.033	-18.51
mother smokes	-0.007	-2.48
mother former smoker	-0.054	-16.81
n.s. friends	0.012	-46.06
smoking friends	-0.004	-12.72
smoking teachers	0.017	5.52
taught smoking risky	0.053	16.24
good student	0.022	4.37
medium student	0.006	1.45
bad student	-0.030	-3.57
has job	0.036	9.14
hours worked / week	-0.001	-11.46
health above avg	0.043	14.73
health below avg	-0.063	-11.47
workplace restrictions	-0.023	-6.85
# smoking risks named	-0.018	-18.60
knows lgl. age & age=19	0.060	8.94
knows legal smoking age	0.080	23.71
male	-0.016	-6.34
constant	0.495	68.64

Note: Table show results of OLS regression of marginal treatment effects ($MTE(X)$) on X. Standard errors not corrected for sampling variability in structural parameters.

Table 6
Price elasticities.

smoking decision	elasticity
early initiation (D)	-1.268
late adolescent if early initiator (Y_1)	0.012
late adolescent if not early initiator (Y_0)	-0.165

Notes: Table gives elasticities of smoking participation to contemporaneous price changes. Calculated as arc elasticities from structural model in response to a 10% price increase.

Table 7
Policy simulation:
Prices held constant at 1990 levels, 1990-1994

	Observed	Predicted (historical prices)	Simulated (counterfactual prices)
early initiation	0.130	0.130	0.160
smoking in late adolescence	0.230	0.228	0.242

Notes: Table shows observed frequency of early smoking initiation and late adolescent smoking, predicted mean probabilities from the estimated structural model at historical prices, and simulated probabilities if cigarette prices were held constant at September 1990 levels.

Figure 1. *Provincial tobacco prices in Canada over time*
Each point represents real price in one province in one month.

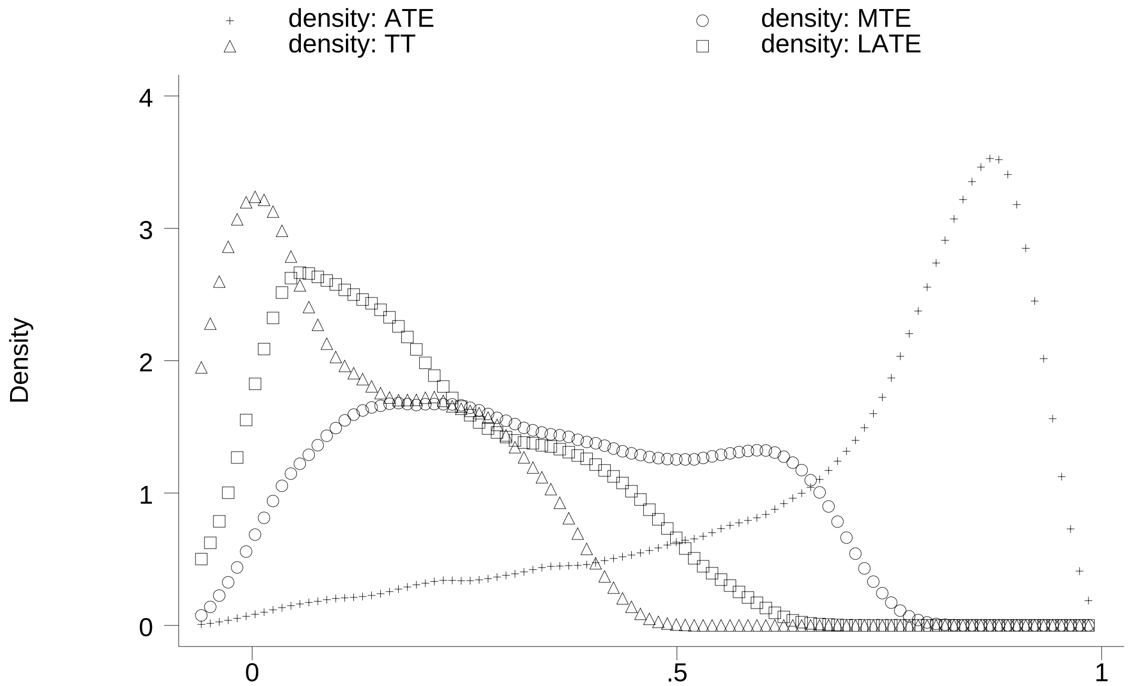

Figure 2. *Heterogeneous addiction estimates.*
Kernel density estimates of distributions of causal effect of early initiation on late adolescent smoking probability calculated from estimated structural model.